Integrating Metadata Schema Registries with Digital Preservation Systems to Support Interoperability: a Proposal

Michael Day
UKOLN, University of Bath, UK
m.day@ukoln.ac.uk

Abstract

There are a large number of metadata standards and initiatives that have relevance to digital preservation, e.g. those designed to support the work of national and research libraries, archives and digitization initiatives. This paper introduces some of these, noting that the developers of some have acknowledged the importance of maintaining or re-using existing metadata. It is argued here that the implementation of metadata registries as part of a digital preservation system may assist repositories in enabling the management and re-use of this metadata and may also help interoperability, namely the exchange of metadata and information packages between repositories.
Keywords: *metadata registries, preservation metadata, digital preservation, OAIS model*

1. Introduction

Ensuring the long-term preservation of information in digital form will be one of the greatest challenges for the information professions in the twenty-first century. While there has been an awareness of digital preservation problems for some time, their importance have recently been magnified because of the increasing dependence of the world on computers and networks. For example, the recent rapid growth in the use of the Internet in the past decade has demonstrated, how much it has become part of the "the fabric of our lives" [1]. Chen has described the digital preservation problem as the "critical, cumulative weakness in our information infrastructure" [2].

Since the mid-1990s there has been a growing awareness of the part that metadata can play in supporting the long-term preservation of digital objects. The preservation function is integral to some definitions of metadata. For example, Cunningham defines it as "structured information that describes and/or allows us to find, manage, control, understand or preserve other information over time" [3]. Calanag, Tabata and Sugimoto have commented that extensive metadata is the best way of minimizing the risks of a digital object becoming inaccessible [4].

This paper will introduce a range of preservation metadata initiatives including the influential Open Archival Information System (OAIS) reference model and a number of other initiatives originating from national and research libraries, digitization projects and the archives community. It will then comment on the need for interoperability between these specifications and propose that the implementation of metadata registries as part of a digital preservation system may help repositories manage diverse metadata and facilitate the exchange of metadata or information packages between repositories.

2. Preservation metadata initiatives

Most digital preservation strategies depend upon the capture, creation and maintenance of metadata [5]. Preservation metadata has been defined as all of the various types of data that will allow the re-creation and interpretation of the structure and content of digital data over time [6]. Defined in this way, it is clear that such metadata needs to support a number of distinct, but related, functions. Lynch, for example, has written that within a digital repository, "metadata accompanies and makes reference to each digital object and provides associated descriptive, structural, administrative, rights management, and other kinds of information" [7]. The wide range of functions that preservation metadata will be expected to fulfil means that the definition of standards is not a simple task and that most of the currently published schemas are either extremely complex or only attempt to produce a basic framework that can be implemented in different ways. The situation is complicated further by the perception that different kinds of metadata will be required to support different digital preservation strategies or digital information types.

To date, preservation metadata initiatives have tended to originate in one of three distinct contexts, from national and research libraries, the archives and records management domain and digitization projects. The following sections will briefly introduce some of these.

2.1. National and research libraries

National and research library initiatives have tended to be either a pragmatic response to management needs, e.g. the element sets developed by the national libraries of Australia [8] and New Zealand [9] or metadata frameworks based on the *Reference Model for an Open Archival Information System* (OAIS), e.g. the specifications developed by the Cedars [10] and NEDLIB (Networked European Deposit Library) [11] projects.

The OAIS model [12, 13] is an attempt to provide a high-level framework for the development and comparison of digital archives. It aims to provide a common framework that can be used to help understand archival challenges and defines a high-level common language that can facilitate discussion across the many different communities interested in digital preservation. The model defines both a functional model and an information model. The functional model outlines the range of functions that would need to be undertaken by a repository, and defines in more detail those functions described within the OAIS specification as access, administration, archival storage, data management, ingest and preservation planning. The information model defines the broad types of information (or metadata) that would be required in order to preserve and access the information stored in an OAIS-based system.

The OAIS information model defines a number of different Information Objects that cover the various types of information required for long-term preservation. The basic assumption of the model is that all Information Objects are composed of a Data Object -which for digital data would typically be a sequence of bits - and the Representation Information that would permit the full interpretation of this data into meaningful information. The OAIS model defines four distinct Information Objects:

- *Content Information* - the information that requires preservation
- *Preservation Description Information (PDI)* - any information that will allow the understanding of the Content Information over an indefinite period of time
- *Packaging Information* - the information that binds all components into a specific medium
- *Descriptive Information* - information that helps users to locate and access information of potential interest.

The information model further divides the PDI into four groups, based on the categories of reference, context, provenance and fixity.

The OAIS information model also defines a conceptual structure for Information Packages. These are viewed as containers that logically encapsulate Content Information and its associated PDI within a single Data Object. Information Packages are defined for submission (SIP), archival storage (AIP) and dissemination (DIP). Of these, the Archival Information Package (AIP) is the most important for digital preservation, as it contains, in principle, "all the qualities needed for permanent, or indefinite, Long Term Preservation of a designated Information Object" [12, p. 4-33]. Those preservation metadata initiatives that have been informed by the OAIS information model have, therefore, tended to concentrate on the definition of AIPs, and specifically on the definition of Content Information and PDI.

Both of the metadata specifications developed by the Cedars and NEDLIB projects were broadly structured according to the OAIS information model's definition of an AIP, defining the Representation Information associated with a Content Object and its PDI. These initial attempts to define preservation metadata, together with the NLA specification, were taken forward by a working group convened in 2000 by OCLC Online Computer Library Center and the Research Libraries Group (RLG). This group first generated a state-of-the-art report that provided a comparison and mapping between the NLA, Cedars and NEDLIB element sets [14]. The working group then produced proposals for Content Information and PDI that were collected together and published in June 2002 as *A metadata framework to support the preservation of digital objects* [15].

Like the Cedars and NEDLIB specifications, the OCLC/RLG metadata framework uses the OAIS information model as part of its basic structure. Therefore, the recommendation for Content Information includes the Content Data Object (a bit-stream) and as Representation Information, elements that relate to the object itself (e.g., file descriptions, significant properties) or its hardware and software environment (e.g., operating systems). The Provenance Information is organized according to an event-based model, defining generic elements associated with processes that might be carried out on the Content Digital Object, e.g. transformations undertaken at ingest, format migrations, etc. The working group did not envisage that the whole metadata framework would be utilized for each and every object within a preservation system, but that metadata would be implemented at varying levels of specificity. They noted that the elements were not necessarily atomic and that it was "easy to imagine cases where the needs and characteristics of particular digital archiving systems ... [would] require deconstruction of these elements into still more precise components."

In 2003, a new group called PREMIS (Preservation Metadata: Implementation Strategies) was formed by the same sponsoring organizations to look at the metadata framework and investigate in more detail the practical aspects of implementing preservation metadata in digital preservation systems.

2.2. Recordkeeping metadata initiatives

The archives and records professions have also been investigating what information might be required to support the long-term preservation of digital objects. As might be expected, their primary focus is on records, defined by the ISO (International Organization for Standardization) Records Management standard [16] as "information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business" [17]. Recordkeeping metadata specifications, therefore, tend to have a strong emphasis on the development of systems that ensure the authenticity and integrity of electronic records.

One of the earliest metadata specifications was based on the Business Acceptable Communications (BAC) model developed by the Functional Requirements for Evidence in Recordkeeping project (the Pittsburgh Project). This proposed a metadata structure that would contain a 'handle layer' for basic discovery data while other layers would store information on terms and conditions of use, data structures, provenance, content and the use of the record since its creation [18]. It was envisaged that much of this information would be automatically generated at the time of creation, would be directly linked to each record, and would be able to describe the content and context of the record as well as enabling its decoding for future use [19].

Together with other developments, the Pittsburgh Project inspired a series of new recordkeeping metadata initiatives, especially in Australia. For example, in 1999 the National Archives of Australia (NAA) published a *Recordkeeping Metadata Standard* that defined the metadata that the NAA recommended should be captured by the recordkeeping systems used by Australian government agencies [20]. Another significant development was the development of the Victorian Electronic Records Strategy (VERS) that defined a self-documenting exchange format (the VERS Encapsulated Object) that permitted the transfer of record content (and metadata) over time [21]. At the current time, the VERS Encapsulated Object is implemented as an XML (Extensible Markup Language) object, chosen because this can be interpreted using basic text editing tools.

All of these developments have been informed by the development of a framework known as the Australian Recordkeeping Metadata Schema (RKMS) by a research project led by Monash University. The project, amongst other things, attempted to specify and standardize the whole range of recordkeeping metadata that would be required to manage records in digital environments [22]. The RKMS also was concerned with supporting interoperability with more generic metadata standards like the Dublin Core and relevant resource discovery schemas like the AGLS Metadata Standard. The RKMS defined a highly structured set of metadata elements conforming to a data model based on that developed for the Resource Description Framework (RDF) [23]. The schema was designed to be extensible and to be able to inherit metadata elements from other schemas.

2.3. Digitization initiatives

Some of the first projects and initiatives to consider the need for preservation metadata were those involved in digitization. The considerable financial investment in digitization means that there was a need to consider the metadata requirements for the long-term management of digitized materials [24].

One recent development has been the development of the Metadata Encoding & Transmission Standard (METS). This is an attempt to provide an XML Schema for encoding metadata that will aid the management and exchange of digital library objects [25]. A METS document consists of seven sections: the METS header, descriptive metadata, administrative metadata, file section, structural map, structural links, and behavior - some of which are intended to group together all of the files that make up a particular digital object and to link content and metadata to a particular structure. The administrative metadata section is intended to store technical information about the file, as well as information about intellectual property rights held in the resource, the source material, and provenance metadata that records relationships between files and migrations.

As part of a separate initiative, the US National Information Standards Organization (NISO) has issued a specification of *Technical metadata for digital still images* for review as a draft standard for trial use (Z39.87) [26]. The draft standard is intended to help define a standardized way of recording the technical attributes of digital images and the production techniques associated with them. The data dictionary includes elements that will record detailed information about images themselves (e.g., formats, compression techniques, etc.), the image creation process, quality metrics, and change history (e.g., migrations). No particular encoding of the elements is recommended, although the Network Development and MARC Standards Office of the Library of Congress is maintaining an XML schema implementation of it called MIX [27]

2.4. Other initiatives

In addition, there are many other metadata standards that contain terms that have relevance to digital preservation. Some of these are format specific or intended for use in particular domains. For example, the MPEG-7 standard [28] is intended to support the management of audio-visual content, and its description schemes can store information about compression methods, data size, access conditions, etc. [29]. The IEEE (Institute of Electrical and Electronics Engineers) Learning Object Metadata (LOM) standard includes elements that describe technical requirements and remarks on installation [30]. Even relatively simple metadata, e.g. the Dublin Core elements specified for use by the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) [31], are likely to contain some data that will be of use to preservation systems. It is becoming clear that some of the most important challenges will be making best use of all the relevant metadata that exists in other forms and securing interoperability between the formats used by repositories.

3. Preservation metadata and interoperability

The plethora of metadata standards and formats that have been developed to support the management and preservation of digital objects leaves us with several questions about interoperability. For example, will repositories be able to cope with the wide range of standards and formats that exist? Will they be able to transfer metadata or information packages containing metadata to other repositories? Will they be able to make use of the 'recombinant potential' of existing metadata?

It is assumed that the generation and maintenance of preservation metadata will be expensive, but it remains a prerequisite of ensuring the successful preservation of digital objects. The difficulty of ensuring digital preservation without metadata may mean that it is ultimately a cheaper and more effective option than the alternative. Chen has written, "although more semantics in metadata will increase costs, it will minimize human intervention in accessing data; seamless support, transition of stewardship and lifetime maintenance will improve" [2]. One way of reducing metadata costs might be to re-use or re-purpose, wherever possible, any metadata that already exist. Hedstrom [32] has argued that there is a need to identify which aspects of existing metadata standards could be used (or adapted) to support recordkeeping requirements, and the same principle applies to preservation requirements more generally.

Recognizing that much existing metadata could prove useful for preservation, both the Cedars specification and the OCLC/RLG metadata framework proposed elements or sub-elements that would be able to store this. For example, the Cedars specification [10] defined 'Existing Metadata' and 'Existing Record' sub-elements as part of its Reference Information and noted that it could either accompany the resource on ingest or be discovered later. The specification, however, is less clear on how this information would be managed or used, restricting itself to a comment that it *could* be integrated using a container architecture like that proposed by the Warwick Framework [33]. The OCLC/RLG framework inherited the same sub-elements from the Cedars specification, but the working group suggested that they could be populated either by a pointer (e.g., an URL) to metadata stored elsewhere or to another information package stored within the system itself. The group said, "it may be more appropriate for the archive to package the metadata into an AIP in its own right, which would then be ingested into the archive" [15].

Packaging metadata into a separate AIP would help to preserve its own structure and content, and the repository management system could ensure that the metadata AIP is securely linked with the AIP containing the data object to which it refers. Under the OAIS model, the metadata AIP would need to contain both data object and any metadata associated with it, which in this case would presumably include the relevant metadata specification. So a single MARC 21 record, for example, would be encapsulated (at least conceptually) with copies of the relevant MARC standards (e.g., formats, code-lists, etc) and any other metadata deemed necessary for its interpretation. However, there are several potential problems with this AIP-based approach. Firstly, it seems rather resource-intensive, e.g. multiplying the potential number of AIPs associated with each preserved object within a repository system. Secondly, the nature of metadata means that they are usually grouped together in some way, e.g. in databases, library systems, etc. If these are also candidates for preservation, there is likely to be some duplication as databases and individual records are preserved separately. One way of overcoming some of these difficulties would be for repositories to build (or utilize) registries of the better-known metadata standards.

In this context, it is interesting that several initiatives are already thinking about format registries to support digital preservation. In 2000, Lawrence *et al*. argued that there was "a pressing need to establish reliable, sustained repositories of file format specifications, documentation, and related software" [34]. Several initiatives are now starting to experiment with these. For example, the DSpace repository system includes a 'bitstream format registry,' which is a way of allowing users to precisely identify the format of the resources that they submit to the system [35]. Also, the Mellon Foundation is funding a research project at University of Pennsylvania on the further development of a Typed Object Model (TOM) that permits the specification of different formats to support their interpretation or conversion [36]. It could be argued that metadata registries are just an extension of this basic concept.

A further issue that will need to be addressed is that of interoperability between repositories. Although the precise way in which intra-repository co-operation will work remains to be worked out in detail, it seems likely that repositories will need to exchange information packages or metadata with other repositories. One approach to this problem might be to develop standard 'exchange-formats,' possibly based on existing standards like METS, VERS or the OCLC/RLG framework. In some domains, e.g. within specific professional communities, it is possible that these formats may emerge as part of the ongoing process of co-operation. In other contexts, it is possible that the exchange of information packages between repositories may become dependent on the sophisticated format conversion facilities that could be offered by metadata registries.

4. Metadata registries and preservation

Metadata registries have been defined as "formal systems that can disclose authoritative information about the semantics and structure of the data elements that are included within a particular metadata scheme" [37]. They can include, for example, definitions of terms used, element usage, permitted schemes, and mappings to other standards [38].

4.1. Existing registry types

Existing metadata registries take many forms. Some are directories of the data elements used in databases, e.g. based on the ISO/IEC 11179 series of standards [39]. An example of this type of registry would be the Environmental Data Registry (EDR) provided by the US Environmental Protection Agency [40]. This service provides authoritative information on the definition, source and uses of environmental data. Other types of registry have been designed to support particular types of encoding schemes, formats or subject domains. For example, the XML.org Registry and Repository set up by the Organization for the Advancement of Structured Information Standards (OASIS) provides information on XML Schemas and Document Type Definitions (DTDs), with the aim of minimizing their overlap and duplication [41]. The DCMI (Dublin Core Metadata Initiative) Metadata Registry contains information on Dublin Core terms and the relationships between them with the purpose of promoting "the discovery, reuse and extension of existing semantics, and to facilitate the creation of new vocabularies" [42]. The SMPTE (Society of Motion Picture and Television Engineers) Metadata Registry stores authoritative information on data elements (labels) or specifications used for audiovisual content and can also be used for the reconciliation of other metadata schemes (e.g. MPEG-7) within the SMPTE infrastructure [43, 44]. The MEG (Metadata for Education Group) Registry aims to provide implementers of educational systems with the means of sharing information about their metadata schemas and to support the re-use of existing schemas [45].

Registries can be implemented in several different ways, e.g. as a database or encoded in a structured syntax like RDF [23, 46]. The advantages of using RDF are that schemas can be managed in a distributed way, but linked by the Internet, and that adherence to its model means that the data can be to some extent machine-understandable. Registries based on RDF include the MEG Registry, which developed a schema creation tool to create RDF data for learning resource metadata [45]. The same broad approach has also been taken by the CORES project, which sees registries as a key part of Semantic Web development [47].

4.2. Basic functions

A metadata registry component of a digital preservation system would have several functions. First, like other types of registry, it would act as an authoritative source of information about the metadata terms and vocabularies used within the repository. Wherever possible, metadata would be kept in its original format and the registry would provide information on how it should be interpreted and gives information on its context. The repository can add (or import) information on new metadata schemas when they become available.

Secondly, once the registry has been populated, it can be used to support the ingest process by providing mappings that could be used to help populate the metadata used by the repository itself. So, for example, a repository could receive an object with its metadata encoded (say) in a particular version of METS. The registry would maintain or give access to the relevant METS specifications (to provide context) and map its terms to the schemes used in the repository. In some cases, there may be multiple instances of metadata accompanying each ingested object, e.g. descriptive information from a MARC 21 record, technical metadata from NISO Z39.87, and rights management information from the Indecs Framework [48] or MPEG-21 [49]. Assuming that the registry maintains mappings from all versions of these standards, the repository could help automatically populate the metadata it requires for managing the data and for generating AIPs.

Thirdly, the mappings maintained within the registry could work in the other direction, to help support the export of metadata or information packages from the repository. Metadata could be collected from the administrative part of the repository and from AIPs, and the registry used to automatically generate selected export formats. While it is highly unlikely that there will ever be a single preservation metadata standard that will be able to be used by all repositories, it may be possible for the different communities to move towards the definition of some kind of standard that might facilitate the exchange of metadata and information packages between repositories.

Figure 1. is an attempt to map these registry functions on to the ingest, archival storage and access parts of the OAIS functional model [12]. In practice, it is likely that other parts of an OAIS, e.g. the data management and administration functions, would also make use of information stored in the registry.

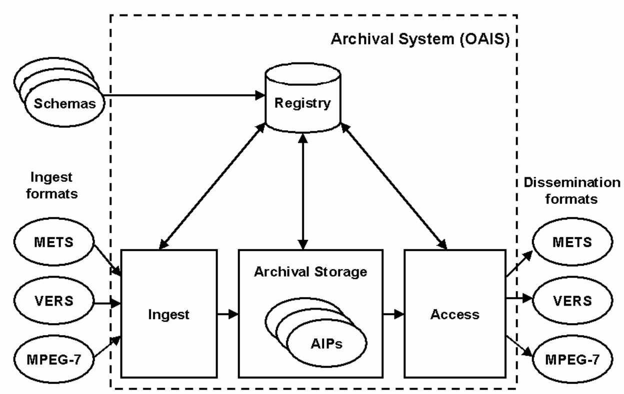

Figure 1. Simplified view of registry functions

4.3. Potential challenges

We have argued that metadata registries could be a useful part of a digital repository system. However, they are definitely not a solution to all problems and many challenges remain. For example, the likely existence of multiple metadata instances about digital objects will mean that the registry may have to make qualitative decisions when there are differences between them. Secondly, any attempt to link metadata will depend upon the universal deployment of unique identifiers; otherwise it will be difficult or impossible for the system to know exactly which metadata should be combined. Also, the registry itself will need to be maintained by the repository and migrated to new systems when necessary. As the number of terms and standards included grows, this task will become increasingly more complicated.

Many current metadata registries are based on Semantic Web technologies like RDF. We have already noted that one of the advantages of RDF is that the schema definitions can be maintained separately from the registry itself. This may have some advantages for repositories, especially where they are working within a global network of 'trusted repositories.' Either standards organizations or repositories can take responsibility for maintaining particular RDF-based schemas and their mappings to other standards. Each individual repository can then select which of these schemas their registry will need to support. A drawback of this approach is that it would require extensive co-operation between repositories and it could become problematic if the schemas themselves became unavailable for whatever reason.

In the longer term, it may be possible for preservation systems to utilize schema registries developed for other purposes. This would be particularly viable where authoritative registries are maintained for particular formats or standards. Naturally, the process could work in the other direction, and the information stored in preservation registries could be made more widely available, e.g. for use by metadata implementers. The maintenance of metadata registries is likely to be a highly resource intensive activity, so the sharing of data between registries may be one way of potentially reducing costs. Progress in this area will be, however, dependent upon some level of standardization, whether based on RDF, XML, or other technologies.

5. Conclusions

A wide range of metadata standards have been developed that have relevance to digital preservation. This paper has argued that metadata registries may be a useful way of helping to manage this diverse metadata within a digital preservation system, and to preserve aspects of its context and original functionality. Registries could also contain authoritative mappings between different standards, thereby helping to facilitate the exchange of metadata or information packages between repositories and end users.

A great deal of work needs to be done before this registry-based approach can be proved to be useful. While it would undoubtedly be useful to have registries of the main metadata standards developed to support preservation, it is less clear how mapping-based conversions between them would work in practice. Metadata specifications are based on a range of different models and conversions often lead to data loss. Also, much more consideration needs to be given to the practical issues of implementation.

References

[1] Castells, M. (2001). *The Internet galaxy: reflections on the Internet, business, and society*. Oxford: Oxford University Press.

[2] Chen, S.S. (2001). The paradox of digital preservation. *Computer*, 34 (3), 24-28.

[3] Cunningham, A. (2000). Dynamic descriptions: recent developments in standards for archival description and metadata. *Canadian Journal of Information and Library Science*, 25(4), 3-17.

[4] Calanag, M.L., Sugimoto, S., & Tabata, K. (2001). A metadata approach to digital preservation. In: *Proceedings of the International Conference on Dublin Core and Metadata Applications 2001* (pp. 143-150). Tokyo: National Institute of Informatics. Retrieved July 21, 2003, from http://www.nii.ac.jp/dc2001/ proceedings/product/paper-24.pdf

[5] Day, M. (2001). Metadata for digital preservation: a review of recent developments. In: *Research and Advanced Technology for Digital Libraries: 5th European Conference, ECDL 2001* (pp. 161-172). Lecture Notes in Computer Science, 2163. Heidelberg: Springer-Verlag.

[6] Ludäsher, B., Marciano, R., & Moore, R. (2001). Preservation of digital data with self-validating, self-instantiating knowledge-based archives [Electronic version]. *SIGMOD Record*, 30(3): 54-63.

[7] Lynch, C. (1999). Canonicalization: a fundamental tool to facilitate preservation and management of digital information. *D-Lib Magazine*, 5(9), September. Retrieved July 21, 2003, from http://www.dlib.org/ dlib/september99/09lynch.html

[8] Phillips, M., Woodyard, D., Bradley, K., & Webb, C. (1999). *Preservation metadata for digital collections: exposure draft*. Retrieved July 21, 2003, from the National Library of Australia Web site: http:// www.nla.gov.au/preserve/pmeta.html

[9] National Library of New Zealand. (2002). *Metadata standards framework -- preservation metadata*. Retrieved July 21, 2003, from http:// www.natlib.govt.nz/files/4initiatives_metaschema.pdf

[10] Russell, K., Sergeant, D., Stone, A., Weinberger, E., & Day, M. (2000). *Metadata for digital preservation: the Cedars project outline specification*. Retrieved July 21, 2003, from the University of Leeds Web site: http://www.leeds.ac.uk/cedars/metadata.html

[11] Masanès, J., & Lupovici, C. (2001). Preservation metadata: the NEDLIB's proposal. *Zeitschrift für Bibliothekswesen und Bibliographie*, 48(3-4), 194-199.

[12] CCSDS 650.0-B-1. (2002). Reference model for an Open Archival Information System (OAIS). Retrieved July 21, 2003, from the Consultative Committee for Space Data Systems Web site: http:// wwwclassic.ccsds.org/documents/pdf/CCSDS-650.0-B-1.pdf

[13] ISO 14721:2003. Space data and information transfer systems -- Open archival information system -- Reference

model. Geneva: International Organization for Standardization.
[14] OCLC/RLG Working Group on Preservation Metadata. (2001). *Preservation metadata for digital objects: a review of the state of the art.* Retrieved July 21, 2003, from the OCLC Web site: http://www.oclc.org/research/pmwg/presmeta_wp.pdf
[15] OCLC/RLG Working Group on Preservation Metadata. (2002). *A metadata framework to support the preservation of digital objects*. Retrieved July 21, 2003, from the OCLC Web site: http://www.oclc.org/research/pmwg/pm_framework.pdf
[16] ISO 15489:2002. Information and documentation -- Records management. Geneva: International Organization for Standardization.
[17] Healy, S. (2001). ISO 15489 Records Management - its development and significance. *Records Management Journal*, 11 (3), 133-142.
[18] Bearman, D., & Sochats, K. (1996) *Metadata requirements for evidence*. Retrieved July 21, 2003, from the Archives & Museum Informatics Web site: http://www.archimuse.com/papers/nhprc/BACartic .html
[19] Bearman, D., & Duff, W. (1996). Grounding archival description in the functional requirements for evidence. *Archivaria*, 41, 275-303.
[20] National Archives of Australia. (1999). *Recordkeeping metadata standard for commonwealth agencies*, v. 1.0. Retrieved July 21, 2003, from http://www.naa.gov.au/recordkeeping/control/rkms/summary.htm
[21] Public Record Office Victoria. (2000). *PROS 99/007 Standard for the management of electronic records*, v. 1.2. Retrieved July 21, 2003, from http://www.prov.vic.gov.au/vers/standards/pros9907.htm
[22] McKemmish, S., Acland, G., Ward, N., & Reed, B. (1999). Describing records in context in the continuum: the Australian Recordkeeping Metadata Schema. *Archivaria*, 48, 3-43.
[23] Lassila, O., & Swick, R.R. eds. (2000). *Resource Description Framework (RDF) model and syntax specification*. W3C Recommendation, 22 February. Retrieved July 21, 2003, from the World Wide Web Consortium Web site: http://www.w3.org/TR/REC-rdf-syntax/
[24] Kenney, A. R., & Rieger, O. Y., eds. (2000). *Moving theory into practice: digital imaging for libraries and archives.* Mountain View, Calif.: Research Libraries Group.
[25] METS - Metadata Encoding & Transmission Standard. Retrieved July 21, 2003, from the Library of Congress Web site: http://www.loc.gov/standards/mets/
[26] NISO Z39.87-2002 AIIM 20-2002. Data dictionary - Technical metadata for digital still images. Retrieved July 21, 2003, from the National Information Standards Organization Web site: http://www.niso.org/standards/resources/Z39_87_trial_use.pdf
[27] MIX - NISO Metadata for Images in XML Schema. Retrieved July 21, 2003, from the Library of Congress Web site: http://www.loc.gov/standards/mix/
[28] ISO/IEC 15938:2002. Information technology -- Multimedia content description interface. Geneva: International Organization for Standardization.
[29] Chang, S.-F, Sikora, T., Puri, A. (2001). Overview of the MPEG-7 standard, *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6), 688-695.
[30] IEEE Std 1484.12.1-2002. IEEE Standard for Learning Object Metadata. New York: Institute of Electrical and Electronics Engineers.
[31] Van de Sompel, H. & Lagoze, C. (2002). Notes from the interoperability front: a progress report on the Open Archives Initiative. In: *Research and Advanced Technology for Digital Libraries: 6th European Conference, ECDL 2002* (pp. 144-157). Lecture Notes in Computer Science, 2458. Heidelberg: Springer-Verlag.
[32] Hedstrom, M. (2001). Recordkeeping metadata: presenting the results of a working meeting. *Archival Science*, 1(3), 243-251.
[33] Lagoze, C. (1996). The Warwick Framework: a container architecture for diverse sets of metadata. *D-Lib Magazine*, July/August. Retrieved July 21, 2003, from http://www.dlib.org/dlib/july96/lagoze/07lagoze .html
[34] Lawrence, G.W., Kehoe, W.R., Rieger, O.Y., & Walters, W.H. (2000). *Risk management of digital information: a file format investigation.* Washington, D.C.: Council on Library and Information Resources. Retrieved July 21, 2003, from http://www.clir.org/pubs/abstract/pub93abst.html
[35] Bass, M.J., Stuve, D., Tansley, R., Branschofsky, M., Breton, P., Carmichael, P., Cattey, B., Chudnov, D., & Ng, J. (2002). *DSpace -- Internal Reference Specification: functionality*. Retrieved July 21, 2003, from the DSpace Web site: http://dspace.org/ technology/functionality.pdf
[36] The Typed Object Model (TOM). Retrieved July 21, 2003, from the University of Pennsylvania Web site: http://tom.library.upenn.edu/
[37] Heery, R., Gardner, T., Day, M., & Patel, M. (2000). *DESIRE metadata registry framework*. Retrieved July 21, 2003, from the DESIRE project Web site: http://www.desire.org/html/research/deliverables/D3.5/
[38] Baker, T., Dekkers, M., Heery, R., Patel, M., & Salokhe, G. (2001). What terms does your metadata use? Application profiles as machine-understandable narratives. *Journal of Digital Information*, 2(2), November. Retrieved July 21, 2003, from http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Baker/
[39] ISO/IEC 11179-3:2003. Information technology -- Metadata registries (MDR) -- Part 3: Registry metamodel and basic attributes.
[40] Environmental Data Registry. Retrieved July 21, 2003, from the Environmental Protection Agency Web site: http://www.epa.gov/edr/
[41] XML.org. Retrieved July 21, 2003, from the OASIS Web site: http://www.xml.org/

[42] DCMI Registry Working Group. (2002). *The Dublin Core Metadata Registry*. Retrieved July 21, 2003, from the Dublin Core Metadata Initiative Web site: http://dublincore.org/dcregistry/index.html
[43] Morgan, O. (1998). Wrappers and metadata. *EBU Technical Review*, 277. Retrieved from the European Broadcasting Union Web site: http://www.ebu.ch/ trev_277-morgan.pdf
[44] Morgan, O. (2003). Metadata systems architecture. *SMPTE Motion Imaging Journal*, 112(4), 129-135.
[45] Heery, R., Johnston, P., Beckett, D., & Steer, D. (2002). The MEG Registry and SCART: complementary tools for creation, discovery and re-use of metadata schemas. In: *Proceedings of the International Conference on Dublin Core and Metadata for e-Communities, 2002* (pp. 125-132) Florence: Firenze University Press. Retrieved July 21, 2003, from http://www.bncf.net/dc2002/program/ft/ paper14.pdf
[46] Brickley, D., & Guha, R.V. eds. (2003). *RDF Vocabulary Description Language 1.0: RDF Schema*. W3C Working Draft, 23 January. Retrieved July 21, 2003, from the World Wide Web Consortium Web site: http://www.w3.org/TR/rdf-schema/
[47] CORES Registry. Retrieved July 21, 2003, from the CORES project Web site: http://www.cores-eu.net/ registry/
[48] Indecs Framework. Retrieved July 21, 2003, from the Indecs Web site: http://www.indecs.org/
[49] MPEG-21 Multimedia Framework. Retrieved July 21, 2003, from the Leonardo Chiariglione Web site: http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm